DOCUMENT RESUME

ED 322 292 CE 055 184

AUTHOR Short, Deborah J.; Grognet, Allene G.
TITLE A Review of Workplace Literacy Tests and Testing.
INSTITUTION Center for Applied Linguistics, Washington, D.C.
SPONS AGENCY Hudson Inst., Indianapolis, IN.
PUB DATE Sep 88
NOTE 66p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Adult Basic Education; Adult Literacy; Adults; Basic Skills; Business Correspondence; Education Work Relationship; *Employment Qualifications; *Functional Literacy; Job Skills; Literacy Education; Occupational Clusters; Reading Achievement; Reading Materials; Reading Skills; *Reading Tests; Standardized Tests; *Test Reviews; Videodisks
IDENTIFIERS *Literacy Audits; *Workplace Literacy

ABSTRACT

A review of workplace literacy tests and testing found that the definitions of literacy and, in particular, workplace literacy have changed and are continuing to change. There is no single, widely accepted definition. Descriptions of workplace literacy demands in terms of reading grade levels are inappropriate. Most tests currently in use are inappropriate for the work environment because they: (1) do not contain job-related vocabulary or tasks; (2) are not indicative of the actual demands placed on workers; (3) are of the paper-and-pencil type rather than performance-based; (4) do not contain an oral component; and (5) do not require the examinee to interpret and analyze. An appropriate workplace test should simulate job tasks. A needs analysis of the literacy requirements for each job or job cluster should be undertaken before a workplace literacy test is designed. New technology should be incorporated into any workplace test. Computers and interactive videodiscs permit more flexibility, greater tracking, and better curriculum planning than does any standardized test. Interactive videodiscs allow a wide variety of question types that more closely simulate the workplace. To be more valid, workplace tests should have the following sections: (1) one part, for all occupations, that should address general concerns such as reading indexes, filling out forms, and writing business correspondence; and (2) another part that varies according to occupational cluster and focuses on particular skills required for its cluster. (Includes 31 references and descriptions of 13 tests, providing title, publisher, purpose, audience, type, design, skills assessed, approach/item type, norms, and review comments.) (CML)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED322292

A REVIEW OF WORKPLACE LITERACY TESTS AND TESTING

by

Deborah J. Short
Allene G. Grognet

CENTER FOR APPLIED LINGUISTICS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

G R Bohan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CE 055 184

Prepared for The Hudson Institute
September, 1988

BEST COPY AVAILABLE

THE TESTING OF WORKPLACE LITERACY
EXECUTIVE SUMMARY

This paper is concerned with the assessment of workplace literacy. It discusses various definitions of literacy, the feasibility of using commercial tests to help determine the literacy skills of workers, and the necessity for determining the job literacy requirements for employment positions. It also evaluates the tests most commonly used in the workplace and offers suggestions for composing a workplace test.

FINDINGS

The definitions of literacy and workplace literacy, in particular, have changed over time and are continuing to change.

- There is no single, widely-accepted definition.
- Descriptions of workplace literacy demands in terms of reading grade levels is inappropriate.

Most tests currently in use are inappropriate for the work environment.

- They do not contain job-related vocabulary or tasks.
- Very few are indicative of the actual demands placed upon a worker.
- Most are of the paper and pencil variety, rather than performance-based.
- Very few contain an oral/aural component.
- Many tests do not require the examinee to interpret and analyze.

An appropriate workplace test should simulate job tasks.

- A needs analysis of the actual literacy requirements for each job or job cluster should be undertaken before a workplace literacy test is designed.
- New technology should be incorporated into any workplace test.
  - Computers and interactive videodiscs permit more flexibilty, greater tracking and better curriculum planning than any standardized test.
  - Interactive videodiscs allow a wide variety of question types, ranging from the traditional multiple-choice format with visual stimuli to a manipulative approach which more closely simulates the activities of a workplace.

Workplace tests should have two sections to be more valid.

- Part 1--the same for all professions-- addressing general concerns like reading indexes, filling out forms, writing business correspondence;
- Part 2--varied according to similar occupations (i.e., clerical, service-oriented, machine-related, etc.)--focusing questions to the particular skills required for success in these job clusters.

THE TESTING OF WORKPLACE LITERACY

Workplace literacy has become an important national issue. It is of concern to employees, employers, unions, vocational and adult educators, Congress, and even, consumers. In our increasingly technological society, different workplace demands are being placed on employees. The abilities to use computers, read manuals, follow directions, and communicate information through oral and written modes are more and more crucial to success on the job and to the quality production of goods. *The Written Word* explains, "The estimated yearly cost of illiteracy due to non-productivity, crime and loss of tax revenue is $225 billion" (p 4).

Workers are expected to be more flexible--able to switch companies or jobs within a company, able to learn the new positions quickly and accurately. Frequently, in these instances, new skills must be developed. Companies try to assist their workers by offering training and retraining programs. It is in the course of such programs that literacy problems may become evident. Companies need to be aware of the literacy requirements for both specific occupations and each new task to be performed by the worker. "Ninety-nine percent (99%) of all workers perform some reading related work each day: to keep pace on the job they need to read an average of 113 minutes each day" (*The Written Word,* p 4). Perhaps a worker will be asked to write more reports or use more math computations in such procedures as statistical processing control, or interpret more charts and tables in a new manufacturing process. Training/retraining should be prepared to address needed aspects of literacy: prose, document or quantitative, as well as written and oral communication and critical thinking skills. However, there are two

questions that need to be investigated before this training can be appropriately developed:

1) What are job-specific literacy requirements; and,
2) What are an employee's current literacy skills?

This paper focuses on the second question through the examination of available tests of adult functional literacy and their applicability to the workplace.

Definitions of Literacy

Currently, there are differing definitions of literacy; moreover,definitions of literacy have changed continually throughout the centuries. At one time, literacy was considered the ability to write one's name and address; at another, the ability to write a simple sentence in any language. According to these definitions, very few people would be classified "illiterate" today, and consequently, literacy problems on the job would be minimal. This is not the case, though, and as most literacy researchers would affirm, these definitions are not suitable at present. Even a more commonly accepted determination, which correlates literacy skills with grade level reading proficiency scales (be it 3rd, 4th, or 5th grade), is inappropriate to the current challenges of survival (Harman, 1985; Broussard, 1985). Although literacy "entails technical reading ability, its proper definition must also take cognizance of the language, culture, environment and the uses to which literacy is to be put. Literacy is a contextual ability..." (Harman, 1985, p 13).

Because general definitions of literacy are not applicable standards for measuring the present problems we confront, more specialized terms have

been composed to focus the discussion on particular situations. These terms include "functional literacy," "survival literacy," and "workplace literacy."

Functional or survival literacy definitions are usually based on competencies for daily life that need to be achieved, although Nafziger (1975) notes they are sometimes "defined in terms of a grade level equivalent or some other norm" (p 1). These competencies have been identified by various people and groups according to their perceptions of what people need, primarily by researching the tasks and situations most people encounter daily and determining the skills necessary to complete those tasks. If a person can demonstrate competency in those tasks, then that person can be considered functionally literate.

In a literacy survey of young adults, done for the National Assessment of Educational Progress (NAEP), functional literacy was defined as, "Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential" (Kirsch & Jungblatt, 1986). Another study, the Adult Performance Level (APL) project, defined literacy as having two dimensions: a set of skills interacting with a set of general knowledge areas. These skills included: communication, computation, problem-solving, and interpersonal relations. The five general knowledge areas were: consumer economics, occupational knowledge, government and law, community resources, and health (Northcutt & Hickok, 1979). In 1970, the Division of Adult Education in the US Office of Education made this comment about adult literacy, "The challenge is to foster through every means the ability to read, write, and compute with the functional competence needed for meeting the requirements of adult literacy" (Cervaro, 1981).

Nonetheless, as Valentine (1986) pointed out, although there are many definitions of functional literacy, this fact does not render one, two, or all of them invalid. Rather, a learner, or employer, can choose the definition which has an underlying terminal objective most closely related to that individual's needs. Thus, the requirements to meet that objective are externally imposed by the environment in which that individual must function. The APL project suggested literacy is "closely bound to the technological state of a particular culture. The person who is 'literate' in one culture may be 'illiterate' in another... (A)s technology changes, the requirements for literacy change" (Northcutt & Hickok, 1979, p 186).

This last observation has significance for the workplace as well. Workers on an assembly line, for example, may never have experienced literacy difficulties on the job. The workers were familiar with their job and did the work properly. They performed their reading and computation tasks adequately. Yet, a new environment evolved at the workplace. Automation on the assembly line made their jobs obsolete. To avoid unemployment, they had to accept different jobs where the reading, writing and computing tasks were more difficult. Perhaps, the tasks were coupled with higher order cognitive skills, like analyzing and synthesizing information. Suddenly, these workers, who had performed competently at their old positions, were unable to succeed in their new ones. They had become "illiterate."

The term, workplace literacy, does not have a commonly accepted definition. In general, it refers to the literacy skills required to perform a job successfully, and includes some reading, writing and computing tasks. However, what these skills and tasks entail for each job has not been

delineated. Rush et. al. (1986) define the term 'occupational literacy' more narrowly as "the ability to competently read required, work related materials" (p 1). Although they limit this definition to competence with printed matter, they further explain that job performance is related to linguistic competencies involving writing , listening, and speaking, too.

Literacy Predictions, Estimations, and Deficiencies

As with the definitions of literacy, there is no consensus either as to the number of illiterate or semi-literate adults in the US, despite numerous surveys and censuses conducted since the early 1970's. Some of these studies also sought to determine in which skills employees lacked proficiency in different occupations. Louis Harris and Associates, ALP, NAEP, and the Hudson Institute, among others, have published their figures and predictions. The percentages, as implied earlier, varied widely depending upon which definition of literacy was used, but nonetheless, all surveyors and researchers agree that a literacy problem exists and needs to be addressed.

The following paragraphs present some of the literacy statistics and predictions in previous studies:

a) Louis Harris and Associates (1971) conducted a nationwide survey of 1747 adults age 16 years or older to determine their functional reading ability. They composed a questionnaire with ten features: 1) telephone dialing, 2) classified housing ads, 3) classified employment ads, 4) personal identification, 5) employment, 6) income, 7) housing, 8) automobile, 9) medical, and 10) citizenship. Features 4 - 10 involved an application form.

Their results indicated that four percent (4%) of the surveyed population "suffered from serious deficiencies in functional reading ability. Their failure on more than twenty percent (20%) of the questionnaire places in serious doubt their ability to 'survive' in practical situations where reading ability is essential" (p 57).

b) At a conference sponsored by the American Council of Life Insurance in May 1983, Niebuhr presented statistics showing that 1 in 5 adults lack the basic skills and knowledge to function in society. Approximately 40 million adults are only marginally literate. Thirteen percent (13%) of 17 year olds are functionally illiterate. Delker, at the same conference, cited the NAEP and APL studies which imply various levels of literacy:
1) functionally illiterate; 2) functionally literate, but not proficient; and 3) proficient. He reported that the 1975 APL project determined fifteen percent (15%) of the working population were functionally illiterate and projected twenty-seven percent (27%) would be by 1983.

c) Harman (1985) cites statistics that predict 72 million adults over the age of 16 are either functionally illiterate or only marginally proficient in basic skills. This composite figure can be split into two groups. The first group encompasses approximately 27 million adults who have a very low level of ability. Harman refers to this group as the "hard-core" illiterates. The second, of approximately 45 million adults, includes individuals with a range of reading comprehension and other basic skill problems (p 25).

d) The NAEP survey, conducted with a young adult (21 - 25)

population, consisted of a 30 minute oral interview and a hour session of simulated performance of 100 'real-life' tasks.

NAEP categorized literacy into three distinct scales:

prose literacy: to locate and use information from texts, editorials, news stories, poems, etc.

document literacy: to locate and use information in payroll forms, job applications, schedules, maps, tables, etc.

quantitative literacy: to apply arithmetic operations embedded in print, like balancing checkbooks, tipping, amount of interest in loan advertisements, etc.

The tasks were assigned a numerical position on the scales (based on 500 points) according to their level of difficulty.

NAEP found that only two percent (2%) were judged to have extremely limited literacy skills and consequently, could not participate in the simulated tasks. Ninety-five percent (95%) of that population could perform at a fourth grade reading level. And, although sixty percent (60%) could perform at an 11th grade reading level, only half of them could complete moderately complex tasks (in the 250 - 275 level of the 500 point scale) and less than one-fifth could do the more complex tasks (at the 350 level or above).

On the prose scale, for example, "locating information in a sports article" was at the 250 level and seventy-two percent (72%) of the respondents were able to perform that task successfully. However, "generate an unfamiliar theme from a short poem," (at 375) could only be done by nine percent (9%). On the document scale, "entering personal information of a job application," (at 200) was performed by

ninety-six percent (96%), while "using a bus schedule to select an appropriate bus for given departures and arrivals," (at 350) was only done correctly by twenty percent (20%). Similarly, on the quantitative scale, "totaling a bank deposit entry," (at 225) was completed by ninety-two percent (92%), but "estimating cost using grocery unit price labels," (at 375) could only be done by ten percent (10%) on the young adult population.

The major NAEP conclusion confirms the current belief that adults cannot be readily classified as literate or illiterate. There is a broad range within the literacy scales and all adults fall along the continuum, yet the 'illiterate' end is more heavily weighted than the 'literate' one. Another significant finding explained that minority adults and those with 0 - 8 years of education performed the least well on tasks.

Literacy Skills on the Job

Although these statistics are disparate, the general consensus indicates there is a large number of semi-literate individuals and a strong proportion of them are in the workforce. Despite the differences, these figures are valuable in light of the changing workplace. *Workforce 2000* explained that although forty-two percent (42%) of the current jobs require some post-secondary education, the number will change to fifty-two percent (52%) for the jobs added between the years 1984 and 2000. Twenty-seven percent (27%) of the new jobs would be in the low-skill category, and jobs that are presently considered in a mid-skill level would become the low-skill jobs of the future. The majority of the reading requirements with these jobs would be reading-to-do something and reading-to-access information. Workers would need the ability to

process the information they find. In reading-to-do, though, the workers need only keep the information in the short-term memory. After the action is completed, the information may be forgotten and only the knowledge of where to look up the information retained.

The *Basic Skills in the US Workforce* (1983) report provided the following statistics:

30% of the secretaries had difficulty reading at their job level;

50% of the managers and supervisors were unable to write free of mechanical errors;

50% of the skilled and semi-skilled employees (like bookkeepers) were unable to use decimals and fractions in math problems.

Although these results would rarely force an employee to lose a position, they do have other repercussions. First, the presence of these skill deficiencies would inhibit job promotion. Second, these problems can be costly to companies. For example, Mikulecky et. al. (1987) reported that Mutual of New York must correct or retype approximately seventy percent (70%) of its correspondence. Furthermore, Westinghouse Electric Corporation's defense gear plant in Sunnyvale, California has been involved in court cases because workers were unable to read safety warnings or follow written directions.

Managers, directors and other employers are concerned with maintaining the quality of America's workmanship and ensuring America's ability to compete successfully in the marketplace. They recognize the need to hire capable employees. A Department of Education survey (see *The Bottom Line,* 1988) of 101 small to medium-sized business executives showed they wanted better skilled workers. The executives identified the basic

skills as the abilities: to read and write, to compute, to communicate and to problem solve. Yet, as other surveys have indicated, the pool of qualified workers is diminishing.

Workplace Literacy Testing

In seeking qualified workers, many companies use a variety of measures when they screen, hire and promote employees. Since no commercially available instruments have been written to assess workplace literacy, companies implement their own tesing procedures. The advent of two court cases, *Griggs v. Duke Power (1971)* and *Albemarle Paper Co. v. Moody (1975),* however, resulted in companies guarding their personnel assessments closely. Those cases determined that tests of general reading ability were unsuitable as literacy screening measures unless they accurately reflect the job requirements.

According to the subsequent Uniform Guidelines on Employee Selection Procedures (1978) literacy tests for hiring and promotion had to be nondiscriminatory (see Mikulecky and Diehl, 1979). If they were discriminatory, there were four options. The tests were:

1) no longer used;
2) changed in order to be nondiscriminatory;
3) given a validity study to prove the test assessed skills necessary for the position and indicated successful job performance; or
4) shown to be valid, but the company was to encouraged to search for another measurement instrument.

The 1979 Bureau of National Affairs report (as cited in Mikulecky and Diehl, 1979) discussed pre-employment testing for hiring and selection.

It stated that many companies use pre-employment psychological tests for nonmanagement jobs to show an applicant's ability to read, write, speak, follow directions, and/or perform practical functions. This type of testing tended to be found in larger or nonmanufacturing companies and most often for office/clerical positions. Frequently, though, the companies did not know the grade level equivalent or the readability level of the tests. Mikulecky and Diehl noted that few of the tests have been reviewed for readability. Of those that had, most received comments about the difficulty of the test. They "test reading difficulties which are excessively high compared to the potential demands experienced on the blue-collar jobs for which subjects are being tested..." (p 13).

Paul Blocklyn (1988) conducted some similar research about pre-employment testing. He surveyed 300 human resource managers and, with a forty-seven percent (47%) response rate, concluded:

55% use job-simulated tasks;
24% use typing, word-processing, clerical tests;
5% use writing or essay tests;
5% use math/computation tests;
5% use physical strength tests, such as lifting, back strength and flexibility measurements;
3% use some equipment oriented tasks, like welding or carpentry;
2% use assessment centers for job simulation exercises.

Also used occasionally are psychological, personality and stress tests.

63% of the companies keep all the data confidential.

At present then, companies do use a variety of testing procedures. These tests include: skill tests, such as timed typing exercises; intelligence/ psychological tests; internally developed (or commercial) reading and writing tests; and some more job-specific tests based on training manuals and job-related materials. The trend indicates practical, job-related skills assessments are becoming more common in pre-employment selection.

A corollary to studies of pre-employment testing are studies that have been undertaken to examine the basic skill deficiencies already found in the workplace. By being cognizant of the basic skill problems that exist, employers can search for better assessment measures during the pre-employment stage. Henry (1983) discussed the aforementioned Center for Public Resources report: *Basic Skills in the US Work Force*. Over fifty percent (50%) of the respondents (who were business, school and union leaders) identified the following deficiencies: writing, especially poor grammar, spelling and punctuation; math, especially decimals and fractions; and listening and speaking, especially following oral instructions and expressing ideas and problems. It is interesting to note, however, that reading deficiencies were not so prominent. This fact calls into question the practice of only giving reading tests as a workplace literacy measure in the pre-employment stage.

Evidently, some workplace literacy tests need to be developed, but they must be more job specific than any test currently available. The different literacy demands of different jobs inhibit the construction of one general test. A single general test would not be an appropriate measure for ascertaining an individual's ability to learn/perform a specific job, if the

test is not task-oriented.

Yet, as *The Bottom Line* notes, "Very little research exists about the relationship of literacy to job performance. Much of what exists is sketchy and based on information obtained from studies conducted in the military..." (p 37). To help accomodate for this gap, more and more studies are being conducted to determine the literacy requirements of certain civilian occupations. With the requirements in hand, better tests can be designed.

Workplace Literacy Requirements

Moe, Rush and Storlie (1980) assessed the specific literacy requirements necessary for success in ten occupations: account clerk, auto mechanic, draftsman, electrician, heating and air conditioning mechanic, industrial maintenance mechanic, licensed practical nurse (LPN), machine tool operator, secretary and welder. They studied the 3 job sites and the 3 vocational training sites for each occupation - reviewing required reading materials, recording oral samples, gathering writing samples, and developing key technical vocabulary lists. They summarized their findings into reading and writing demands and compared requirements on the job with instruction in vocational training.

This study revealed inconsistencies between training and job requirements. In some instances training demands were more rigorous than work demands. In other cases, the reverse was true. On the job, all the reading demands were at a minimum 9th grade reading level, except the welder position which had few reading materials-- mainly single word information. In vocational training programs, the minimum reading level was 8th grade, and that was for the welder position. The minimum level

for draftsman, electrician and account clerk was less demanding in the training than on the job. The reading levels for the LPN and the heating and air conditioning mechanic were more demanding. Only the auto mechanic and machine tool operator matched the job and training requirements exactly.

Regarding the math demands, overall, the job and training requirements corresponded. Every job needed the basic arithmetic skills and eight of the ten needed decimals and fractions, as well. Algebra was unnecessary, yet present, for the account clerk and the secretary training; but necessary, yet missing, for the heating and air conditioning mechanic and the industrial maintenance mechanic. For that latter profession, geometry and trigonometry were also found to be used on the job, but were not in the vocational program.

Heinemann (1979) delved further into the reading and writing skills of secretaries. She developed eight job-related performance based tasks, from interviews and analyses of job manuals and related materials, to diagnose those skills. Her subjects, thirty-nine secretaries in a large New York corporation, exhibited three error patterns. They had difficulty in:

1) following directions;
2) distinguishing capitalization, spelling and verb tense mistakes; and
3) judging relative importance; i.e., determining main ideas from details and identifying significant business-related events.

More than half of the secretaries made punctuation errors and omitted part of a business letter. Heinemann's tasks, which engendered some retraining recommendations, could also be used in pre-employment assessment.

Another study about workplace literacy was conducted by Jacob (1982) at a

Baltimore dairy plant. Through interviews with managers and employees, on-site observations, and collection and analyses of manuals, other written materials and alphanumeric symbols, Jacob was able to determine the types of documents (with their corresponding reading and writing assignments) associated with the various positions within the plant. Most of the documents were forms and required some writing. The next frequently used documents were tables, usually one-dimensional with one or two words and numbers. Jacob concluded that workplace literacy for many of these jobs required the ability to comprehend tables and charts, as well as prose materials. Therefore, a workplace literacy test for the plant should contain such document comprehension tasks.

The most significant research and development of workplace literacy has been carried out in the military. Originally begun in the Air Force, research by Sticht and others on job literacy requirements has spread to other services. As mentioned earlier in this report, this research resulted in classifying types of reading: reading-to-do (to perform an action); reading-to-access (to find information); and reading-to-learn (to study and retain information).

Sticht and Caylor (1972) developed job reading task tests (JRTTs) for three army positions: general vehicle repairman, unit and organizing supply clerk, and cook. They decided to assess the enlisted soldiers' ability to read-to-do, i.e., their ability to read the manual to perform a task. First, Sticht and Caylor designed a classification system to categorize various army manuals. They determined that the manuals contained:

1) tables of contents and indexes - for locating information;
2) standards and specifications - for delineating the rules or tolerances that tasks must conform to;

3) identifications and physical descriptions - for recognizing items through a symbolic representation with an I.D. code or a description of distinguishing physical characteristics;
4) procedural directions - for following instructions step-by-step;
5) procedural check points - for reminding the worker of the task through brief summaries or key words, assuming the worker is already familiar with the task; and,
6) functional descriptions - for explaining the intended purpose of an item or task.

Based on these classifications, Sticht and Caylor developed job reading task tests and gave the JRTTs, as well as standard reading tests, to three groups: recruits, men in the first week of training, and men in the seventh week of training. The results indicated that the JRTTs were valid estimates of general reading ability. The correlation between the two types of tests ranged from .65 to .80. Also significant was the result that men with training did better on the JRTTs than men who had similar general reading ability, but no specific job training.

The procedures Sticht et. al. used in developing the JRTTs have been replicated and refined for several other groups in the armed forces. The tests were adjusted to be job-specific and used information from the most frequent types of job reading material. In addition to examining the reading material, the task test developers (see Sticht & Mikulecky, 1984) also interviewed the job performers at work. The task tests included table and flowchart comprehension and information transfer to further authenicize the demands placed on workers using manuals on the job. Sticht and Mikulecky also developed a literacy training program with

pretests and posttests. The program showed an average gain of 0.7 grade level points on general standardized reading test, yet 2.1 points on the job reading task tests. Sticht and Mikulecky concluded testing and training specifically for what is necessary on the job is important and can be achieved.

Diehl (1978) discussed functional literacy and the recent measures used to assess it. He criticized the APL survey because it found positive correlations between literacy skills and three other variables: income, education and job status. Diehl, as did Cervaro (1981), questioned the generalizability of these results. Diehl did not wish these variables to be perceived as the sole bases for determining literacy.

Diehl also reviewed Project REALISTIC, which was Sticht's primary work with the military throughout the 1970's. He pointed out that Sticht used readability, job performance and reading performance measures to develop his job reading task tests. There had been a low correlation ($r = .30$ to $.40$) between general reading ability and on-the-job performance before job-specific training programs were implemented using the job literacy requirements determined by Sticht. The training was successful in part because the materials were job-related and assessed "reading-to-do" abilities, rather than "reading-to-learn" abilities which were less applicable to the jobs.

Commercially Available Literacy Tests

There are some tests that are commercially available and have been commonly used to assess adult basic skills. Those primarily concerned with reading and writing include: ABLE (Adult Basic Learning Examination),

AIRT (Adult Informal Reading Test), BOLT (Basic Occupational Literacy Test), CASAS (California Adult Student Assessment System), GED (General Education Development), READ (Reading Evaluation Adult Diagnosis), and TABE (Test of Adult Basic Education). A widely used oral test is: BVOPT (Bilingual Vocational Oral Proficiency Test) . There are also two tests with literacy components designed specifically for adults whose first language is not English: BEST (Basic English Skills Test) and HELP (Henderson-Moriarity ESL/Literacy Placement). Two workplace tests are: IRT (Industrial Reading Test) and JEVS (Jewish Vocational and Employment Service).

A test can be used for many purposes: assessment, achievement, diagnosis, placement, and program evaluation. As an assessment measure, a test may be used during or after a training session. It would purport to determine how much information is known/has been learned. A diagnostic test can help identify deficiencies in an examinee's knowledge and thus, guide additional instruction. A placement test is given upon entry to a program, so a learner may be offered an appropriate level of instruction. A test used for program evaluation is beneficial to both the instructors and the learners, because the results indicate problematic areas with the instructional design and implementation.

Instructors frequently consider two main features of a test when they review them for possible use in a program. These features are the validity and reliability factors. Validity is necessary for all tests because it indicates that the test actually measures the ability or knowledge it claims to measure. Reliability is important, so an instructor knows that scores will be consistent from one administration of the test to the next.

Several types of validity are defined according to how a test will be used. Content validity is most important. For a test to be content-valid it must be representative of the content presented during instruction; i.e., it must test what was taught. Face validity refers to the subjective impression the examinee has that the test is valid. It is similar to content validity. Concurrent validity tells the extent to which a new test or test item correlates to a pre-determined standard or another accepted test. Predictive validity indicates the ability of the test scores to predict later performance on the subject(s). A well-known example of a test purporting predictive validity is the SAT (Scholastic Aptitude Test) used as a prerequisite for entrance to many colleges and universities.

Reliabilty measures involve statistical calculations. A test with high reliability (approaching r =1.0, although over r = .60 is considered good) suggests the ranking of examinees' scores after one adminstration of the test will be almost the same after a subsequent administration of either an alternate form or the same test.
When the same test is given on two separate occasions, the measure is referred to as Test-Retest Reliability. When the internal consistency of the questions on the test is at issue, either a Split-Half Reliability measure or a Kuder-Richardson Reliabilty measure (KR-20 or KR-21) may be performed. A Split-Half measure divides the test into two parts (usually by the odd and even question numbers) and correlate the scores from each half. A KR-20 or KR-21 measure correlates one item with every other item on the test. It can provide an average correlation for the test.

Besides considering the validity and reliability features of a test, an instructor will also decide if a norm-referenced test or a criterion-referenced test is more suitable for the program. A norm-referenced test

evaluates students according to a pre-determined standard mean or performance level set by a large sample test population. It would list scores in terms of percentiles and compare results of one examinee with the results of all other examinees; i.e., it gives a ranking. The percentiles are only accurate for that particular test administration. The aforementioned SAT is an example of a norm-referenced test.

A criterion- referenced test evaluates students according to their achievement or performance as it is applied to a cut-off score. A criterion-referenced test is usually used with a small testing population. It lists raw scores, usually the percentage correct (based on 100%). The scores may be used anytime since they depend on the content of the test, not the rankings of the examinees. A written driving license test is an example of a criterion-referenced test.

The brief explanations of the tests which follow include information found in test manuals, articles written about the tests, and personal reviews of some of the tests. Although equal amounts of information were not available for all the tests, the general categories for the reviews below are: test title, publisher, test purpose, test audience, test type and design, skills assessed by the test, testing approach and item type, and norms (including reliability and validity figures). The reviews conclude with an evaluation of the test's applicability to the workplace.

TITLE: ABLE (Adult Basic Learning Exam), Level 1

PUBLISHER: The Psychological Corporation, HBJ

PURPOSE: to measure the basic educational achievement of adults

AUDIENCE: adults, primarily those enrolled in adult education programs

TEST TYPE: WRITTEN

LEVELS: 3, each is based on the number of years of formal education the examinee had; level 1 is the most basic--1 to 4 years of education

FORMS: 2 equivalent forms available

TEST DESIGN: paper and pencil test
all multiple choice responses
Level 1 permits dictation of some subtests

SKILLS ASSESSED: vocabulary, reading comprehension, spelling, number operations and problem solving
(Levels 2 and 3 include a language subtest with grammar, capitalization and punctuation.)

TESTING APPROACH/ITEM TYPE:

Several subtests can be dictated for examinees with low literacy skills

Vocabulary - the sentences and answers can be dictated. The sentences are missing one word.

Reading - not dictated. It includes commonplace pictures and signs as well as short passages. Some passages are presented as modified cloze tests, but multiple choices are given for their completion.

Spelling - examinees choose the correct spelling of a word read aloud, read in a sentence and then read aloud again.

Number operations - first five problems are dictated. The problems in this subtest are arithmetic.

Problem solving - word problems are dictated and also seen in print for all questions. Many of the questions include visuals, like charts, graphs and geometric shapes, to help depict the problem.

NORMS: KR-21 for Level 1, Form E r = .78 to .93 (high reliability)
Correlations with SAT scores are only moderate r = .51 to .59
(except spelling r = .69).

The manual explains the writers composed the test according to the content of adult education programs nationwide and thus claim content validity for those common objectives. The suggestion is made, though, for individual programs to determine the validity as the test applies to their curriculum.

COMMENTS:

Although the administration manual states that this subtest is "...designed to assess the knowledge and understanding of words that are frequently encountered by adults in their work or other daily activities," a review of the words reveals not all are necessarily very common. For example, the items include: riddle, kennel, breathtaking, kindling, dorsal fin, brittle, rein, and

mournful. On the positive side, some work-related terms, such as earn, wages, employer, applicant, and tax exempt, are given too.

Some of the vocabulary in the reading subtest is applicable to work, like "steel shoes," "emergency exit," and "apply within."

The spelling subtest includes items more representative of the words adults use in daily communication than the vocabulary list. These items include: was, after, ready, and names.

The number operations subtest has the first five problems dictated along with instructions for solving. The remaining items have no instructions given--neither oral, nor written.

In the problem-solving subtest, the visuals seem more helpful in finding the solution than the prose. Certain knowledge, such as 16 ounces = 1 pound, is assumed. Many of the problems, such as those about payroll deductions or totalling the number of hours worked, are job-related.

The majority of the questions use simple present and past tenses as well as the active voice. However, the presence of conditional and perfect tenses, embedded questions and passive voice (primarily in the problem-solving subtest) show the test is written above a low literacy level.

This test, which is promoted as an achievement test, was not designed for literacy, nor for a workplace assessment. It does include some items and vocabulary appropriate for a workplace

environment, but only generic situations that deal with payroll, transportation costs and job applications. Success on this test would not adequately determine success on the job, since specific job reading, writing and computation demands are not made upon the examinee. Moreover, the language in the questions has not been systematically simplified for a "literacy-level" examinee.

TITLE: AIRT (Adult Informal Reading Test)

PUBLISHER: University of Missouri - Kansas City

PURPOSE: to assess adults' ability to read stories aloud and to assess their accuracy in reading

AUDIENCE: adults enrolled in literacy programs

TEST TYPE: ORAL

LEVELS: only 1

FORMS: 2 forms available

TEST DESIGN: performance-based

SKILLS ASSESSED: oral reading ability and comprehension, pronunciation, word recognition

TESTING APPROACH/ITEM TYPE:

An introduction is provided for each story as motivation.

Six stories are read aloud and examiners checks oral errors, comprehension and rate of reading. Errors are: substitutions, omissions, additions, help from the examiner,and word ending mistakes. Examiner asks some oral comprehension questions about the stories read by the examinee.

The Mitzel word list (a list of 2000 commonly used words) is also used to test word recognition.

NORMS: The Pearson Product correlation test showed no correlation between the AIRT and the ABLE tests.

No data was available for reliability and validity statistics.

COMMENTS:

Leibert (1973) conducted a study comparing ABLE and AIRT. He concluded that informal reading tests (AIRT) better estimate the instructional level of the student whereas standardized tests (ABLE) are more sensitive to small achievement shifts.

He observed that the AIRT analysis showed that a "large percentage of adults in ABE programs are reasonably literate when reading achievement is defined as accuracy of oral reading and literal recall..." (p 34). Many adults had a broad range of ability upon which they functioned adequately with respect to oral reading comprehension and accuracy. The major difference among examinees was the rate of reading aloud.

Leibert's conclusions indicate that the AIRT would not be an accurate test of workplace literacy. The testing of oral reading is not a valid measurement for most jobs since very few positions require an employee to read aloud at a specific rate and then check the comprehension of what s/he has just read. Furthermore, the material in the AIRT test is not designed to assess workplace vocabulary, nor situational problems.

TITLE: BEST (Basic English Skills Test)

PUBLISHER: Center for Applied Linguistics

PURPOSE: for placement, achievement, diagnosis, and program evaluation

AUDIENCE: students in ABE, refugee and immigration, pre-vocational and ESL programs

TEST TYPE: ORAL and WRITTEN

LEVELS: 1

FORMS: 2, but only Form B is used now

TEST DESIGN: part performance-based with the examiner and examinee face to face, part paper and pencil with multiple choice questions

There are two sections: Oral and Literacy

SKILLS ASSESSED: speaking ability, pronunciation, listening comprehension, reading comprehension and writing

TESTING APPROACH/ITEM TYPE:

This test incorporates survival literacy skills and some job-related tasks.

Oral interview - examinees complete simulated real-life speaking and listening comprehension tasks, often with picture cues, using real money, telling time and demonstrating social conversational ability. The responses may be verbal or gestural.

A pronunciation score is given after the entire section is completed.

This section also has one reading and one writing task to determine if the literacy section should be given.

Literacy section

Reading comprehension - includes survival reading tasks with calendars, labels, want ads, and schedules.

Writing - also contains survival tasks like addressing envelopes, writing checks, filling out applications.

NORMS: The Oral Section and the Writing portion are scored for comprehensibilty, not 100% accuracy. Scores can be interpreted along SPL (student performance level) guidelines.

KR-20 r = .79 to .91 for the Oral Section

r = .90 to .97 for the Literacy Section

COMMENTS:

The BEST test can be considered a literacy assessment instrument. The pre-vocational language and tasks required in the test, however, are too general (and more survival-oriented) to accurately measure workplace literacy.

TITLE: BMCT (Bennett Mechanical Comprehension Test), Form T

PUBLISHER: The Psychological Corporation, HBJ

PURPOSE: to assess mechanical ability and knowledge for understanding physical and mechanical principles; to select, evaluate and promote employees

AUDIENCE: employees and candidates for employment in businesses and industries like construction, machine maintenance, utilities and production.

TEST TYPE: WRITTEN

LEVELS: 1

FORMS: 1

TEST DESIGN: paper and pencil with multiple choice responses

(If the examinees have difficulty reading, the questions may be read aloud or tape recorded.)

SKILLS ASSESSED: knowledge of physical and mechanical principles

TESTING APPROACH/ITEM TYPE:

Examinees read or listen to the questions and look at the pictures. They choose point A or B (or sometimes C) located somewhere on the picture to answer the question. An example question is "Where should you put the rock so it will work best as a fulcrum?"

If there is no Point C, than C may be chosen as the default option, e.g., if both A and B are equal, or if neither A, nor B are correct.

Pictures are appropriate to the workplace. There are pulleys, wheels, gears, water pressure items, inclines, et. al.

NORMS: No data was available for reliability and validity values.

COMMENTS:

This is a test of mechanical knowledge, and some common sense, but it is not a language assessment. The questions are not simplified and they utilize perfect and conditional tenses.

The test does measure knowledge appropriate for certain businesses and industries, like construction, production, and machine shops.

TITLE: BOLT (Basic Occupational Literacy Test)

PUBLISHER: US Department of Labor

PURPOSE: to measure basic reading and arithmetic skills
to measure literacy achievement in areas of reading and arithmetic according to occupational requirements

AUDIENCE: educationally disadvantaged adults; test is primarily for State Employment Security Agencies

TEST TYPE: WRITTEN

LEVELS: 4--advanced, high intermediate, basic intermediate, and fundamental (Arithmetic reasoning subtest only has 3 levels--advanced, intermediate and fundamental)

FORMS: 3--A, B, C

TEST DESIGN: paper and pencil
all multiple choice responses
Designed to cover a range of difficulty from a barely literate level to a high school dropout level
Has a brief screening test to determine which level is appropriate for the examinee

SKILLS ASSESSED: Reading vocabulary, reading comprehension, arithmetic computation and arithmetic reasoning

TESTING APPROACH/ITEM TYPE:

All multiple choice answers have 5 possible responses.

Subtests:

Vocabulary - examinees choose the word that best completes a given sentence

Comprehension - examinees read short passages and answer questions.

Computation - examinees solve arithmetic problems given in symbolic form, such as 1.2 x 6 = _.

Reasoning - examinees solve word problems which do not have accompanying visuals

NORMS: for all forms and levels of reading vocabulary and comprehension -
KR-20 r = .61 to .80.
for all forms and levels of arithmetic computation and reasoning -
KR-20 r = .68 to .82.

Uses GED ratings of occupations in the Dictionary of Occupational Titles to interpret BOLT scores by relating the BOLT scores to SAT scores and then to GED conversions.

COMMENTS:

The Fundamental level is written for examinees with low literacy skills, but the upper limit of the range of difficulty, the "High School dropout level," is not clarified. The manual does not explain which grade it refers to--8th grade or 10th grade or what? The content is not applicable to the workplace since it assesses neither job-related vocabulary, nor job-related skills.

This test is not readily available to most businesses and industries. It is used primarily in the selection of government employees.

TITLE: BVOPT (Basic Vocational Oral Proficiency Test)

PUBLISHER: Melton Peninsula, Inc. (Dallas, TX)

PURPOSE: to measure speaking and listening proficiencies, to screen for enrollment into bilingual vocational training programs, to assess achievement of English proficiency during and after training

AUDIENCE: bilingual adults (Spanish/English) enrolled in or entering vocational training programs

TEST TYPE: ORAL

LEVELS: 1

FORMS: 2 (A = pretest, B = posttest)

TEST DESIGN: performance-based

not designed to test discrete grammar points

four subtests require oral or physical responses

SKILLS ASSESSED: listening comprehension, speaking

TESTING APPROACH/ITEM TYPE:

Four subtests include:

Questions and Answers - questions become progressively more difficult;

Open interview - using color posters to obtain a 4 - 5 minute speech sample;

Repetitions - students repeat sentences that increase in length and grammatical difficulty;

Following imperative directions with a physical response - students manipulate objects according to commands.

Content of BVOPT is derived from observations and recordings of language used in vocational training classes.

NORMS: The manual claims BVOPT is criterion-referenced, but the cutoff scores for entry and exit of a vocational training program is not given. The manual only interprets scores for three levels of beginning English proficiency.

No reliability data was available.

Content validity is claimed because the test was based on actual language used in bilingual vocational training classes.

COMMENTS:

This test has no reading or writing subtests, so it cannot be considered a test of literacy. Since the content was derived from actual training classes, it does refer to job-related items, directions and safety precautions.

Messerschmitt (1987), in her review of this test, considers it useful for low-level adults if literacy is not a concern. She does point out, however, that for multilevel classes, the test's utility would be diminshed since no cut-off limits are given to guide placement. She also questions the third subtest of sentence repetition, because it could be testing memory instead of language.

TITLE: CASAS (California Adult Student Assessment System)/GAIN Appraisal Program

PUBLISHER: State of California, Department of Education

PURPOSE: to assess the basic reading, math and functional listening comprehension skills of welfare recipients preparing for employment through items in the CASAS item bank. These items address competencies in 5 general knowledge areas: consumer economics, occupational knowledge, government and law, community resources, and health.

AUDIENCE: adult welfare recipients participating in the GAIN program

TEST TYPE: WRITTEN

LEVELS: 1

FORMS: 1

TEST DESIGN: paper and pencil test

all multiple choice responses

SKILLS ASSESSED: Listening: for limited English proficient participants of functional listening skills

Reading: basic reading comprehension in a functional context

Math: basic math skills and applications to a functional context

TESTING APPROACH/ITEM TYPE:

The reading and math questions have written cues. Many are visual--signs from a workplace, want ads, job application forms, pay stubs, time sheets, dials on meters--or job-related prose--a letter of interest in a position, safety rules, catalogue price listings, course descriptions.

The math involves basic facts and computations.

The answers are all multiple-choice, but the choices may be graphic or written.

NORMS: Based on the Field Test (conducted 7/86 - 12/86)

Reading: KR-20 r=.89, KR-21 r=.88

Math: KR-20 r=.86, KR-21 r=.84

No information was directly available about the validity.

COMMENTS:

The tests are functionally based and contains items related to the workplace. Some of the items are easier and could be used to test literacy skills (especially those with graphics), yet overall, the test does not assess literacy.

The tests, however, are not performance-based.

TITLE: GED (General Education Development)

PUBLISHER: GED Testing Service of the American Council on Education

PURPOSE: to assess adults' knowledge of general outcomes and skills associated with 4 years of regular high school instruction.

AUDIENCE: adults seeking a high school equivalency diploma

TEST TYPE: WRITTEN

LEVELS: 1

FORMS: 1, but available in English, French, Spanish, Braille large print versions, and on audio tapes

TEST DESIGN: paper and pencil test

5 separate tests

all multiple choice responses

SKILLS ASSESSED: writing, reading comprehension, knowledge of social studies, math and science

TESTING APPROACH/ITEM TYPE:

Questions all have written cues and multiple choice responses. The questions do not measure specific details of individual high school courses.

Writing - assesses spelling, capitalization, usage, sentence corrections, logic and organization.

Reading - has comprehension questions based on practical and general reading passages, prose literature, drama and poetry.

Math - contains arithmetic, algebraic and geometric problems.

Science - has sections for biology, earth science, chemistry, and physics.

Social Studies - includes passages and questions about economics, geography, political science, history and behavioral science.

NORMS: KR-20 r = .84 to .90 (high reliability)

The GED score scales are based on a sampling of scores by graduating high school seniors.

Because the content for the GED tests is derived from a sampling of U.S. high school curricula, the tests have been assumed to be valid for more than forty years, with all states accepting the GED equivalency diploma.

COMMENTS:

The GED assumes the examinee is literate and can perform reading, writing and mathematical tasks on a high school level. Thus, the GED is not a literacy test. Furthermore, its content is not oriented to the workplace.

The GED has been considered a standard in Adult Basic Education programs for over forty years. Since the 1970's, though, some programs have switched from using GED requirements to guide their curriculum to using the APL competencies. Cervaro (1981) was skeptical about the validity of the APL competencies and decided to compare the two examinations. The main question raised about the APL concerned its content validity. Gathering samples of all kinds of behavior that lead to functional

competence is extremely difficult, and so, the competency behavior assessed in the APL are a result of the APL developers' judgments. Nonetheless, when Cervaro examined total scores on the APL survey and the GED test statistically, he found a high positive correlation ($r = .81$, $p < .00001$). He subsequently concluded that both would thus be acceptable for ABE curricula, and the selection of one course or another could then be determined by the individual learner's needs. Those who desired a more academic framework or college preparation could follow a GED-based program. Those wanting a more practical course could enroll in an APL-based course.

TITLE: HELP (Henderson - Moriarity ESL/Literacy Placement)

PUBLISHER: Alemany Press

PURPOSE: to determine ESL level placement of adult SE Asian refugees

AUDIENCE: Southeast Asian refugees in ESL classes with:
a) no reading or writing skills;
b) minumum reading and writing skills in native language; or,
c) reading and writing skills with a non-Roman alphabet.

TEST TYPE: ORAL

LEVELS: 1

FORMS: 1

TEST DESIGN: performance-based

SKILLS ASSESSED: speaking, writing

TESTING APPROACH/ITEM TYPE:

There are three sections:

Background information interview and native language assessment;

Oral English assessment - students see picture cues to elicit oral responses, tell time or manipulate money, with some word copying and sight word recognition exercises;

Written English assessment - students fill out a brief form.

The focus is on meaningful communication, not 100% accuracy. A fluency score is given for the first section, communication scores are given for the last two.

NORMS: The test is criterion-referenced.

No data is available for reliability or validity values.

COMMENTS:

An positve feature of HELP is the possibility for the examiner to assist the student in the oral section by providing alternate cues or items if the original question is not understood.

This test does assess some literacy skills, but the language and tasks are not oriented for the workplace. At best, it has a functional purpose with exercises requiring the ability to use a phone and fill out forms.

TITLE: IRT (Industrial Reading Test)

PUBLISHER: The Psychological Corporation, HBJ

PURPOSE: to measure reading ability of written technical materials for vocational schools and industries

AUDIENCE: students in vocational education programs or workers with high school education or equivalency diplomas

TEST TYPE: WRITTEN

LEVELS: 1

FORMS: 2; A - for business and industry, B - for schools and businesses

TEST DESIGN: paper and pencil

multiple choice responses

SKILLS ASSESSED: reading comprehension

TESTING APPROACH/ITEM TYPE:

Students read 9 passages written to be consistent with training manuals and company materials. They include: safety regulations, manual explanations, machine descriptions, memoranda, etc.
Students answer multiple choice questions that ask for main points, supporting details, titles, and inferences.

Designers examined the Occupational Outlook Handbook, the Dictionary of Occupational Titles, job descriptions, education and training requirements and other related materials while developing this test.

NORMS: Testing manual gives percentile ranks based on raw scores, with means and standard deviations.

split-half reliability coefficients: A r = .80 to .90
B r = .79 to .92

alternate form reliability: r = .75

Manual claims the content is valid because the reading passages based on authentic materials.

Correlations with other tests are also given:

BMCT r = .67

BOLT r = .56 reading vocabulary
r = .57 reading comprehension

COMMENTS:

The manual explains that this test is not for use below a 9th grade reading level. Thus, it is inappropriate for literacy level adults.

Although the manual asserts content validity, it does not use authentic materials.

The manual states that good performance on the test does not depend on subject matter knowledge. Since it is primarily a test of reading skill, as well as a paper and pencil test, it may not be suited for a workplace environment where more physical manipulation is required.

TITLE: JEVS (Jewish Employment and Vocational Service)

PUBLISHER: Vocational Research Institute, JEVS Inc.

PURPOSE: to measure an individual's ability to perform workplace tasks by watching a demonstration and following oral instructions

AUDIENCE: primarily vocational education students and adults using the employment service

TEST TYPE: ORAL (and visual)

LEVELS: 1

FORMS: 1

TEST DESIGN: performance-based

SKILLS ASSESSED: listening comprehension, ability to copy demonstrations

TESTING APPROACH/ITEM TYPE:

Examinees watch a demonstrated activity and then perform the task themselves. Instructions are given orally. Samples are left to look at, but examinees should not take them apart.

Tasks include: counting, sorting, assembling, stamping and packaging. Examinees use some tools like rulers, scissors, and needles.

Tasks increase in difficulty, either with the time alloted to complete the task or the type of task involved. Examinees are allowed to ask for some help, though most requests will lower their quality rating.

The evaluator keeps a work sample observation record.

Each task is rated on a 3 point scale based on two criteria: the amount of time taken to complete the task, and the quality of the completion. The quality rating considers neatness, number of errors made, frequency and type of help requested.

NORMS: The scoring is based on percentile ranks that were normed at the JEVS Philadelphia Center, but developers encourage other centers to set their own norms and offer to help them do so.

COMMENTS:

This is a workplace test since activities are similar to those found in various businesses and industries. It corresponds more closely than other tests to what actually happens at a workplace, especially with the demonstrations and oral directions. It is good for low level literacy adults since very little reading and no writing are involved. By the same token, however, since those skills are not assessed, this test cannot be regarded as a literacy measurement.

Examiness need some basic abilities, such as, counting, using scissors, and reading diagrams with a few specialized words like "joint" and with inch symbols. They must also not be colorblind.

A unique feature of the JEVS test is the permissibilty of requesting help. On the job, employees would ordinarily ask for assistance if necessary. The evaluators of this test will assist the examinees to a certain extent, although the type and amount of assistance needed for a particular task may count against them in the rating.

TITLE: READ (Reading Evaluation - Adult Diagnosis)

PUBLISHER: Literacy Volunteers of America, Inc.

PURPOSE: to assess the reading needs and progress of adult students

AUDIENCE: adults in literacy programs

TEST TYPE: ORAL and WRITTEN
LEVELS: 1
FORMS: 2--for parts 1 and 3

TEST DESIGN: There is a mixture of oral and written cues.
Examiners show students the testing booklet with the cues and record their responses.

SKILLS ASSESSED: oral reading, sight word recognition, word analysis, reading (or listening, if examiner reads to the student) comprehension

TESTING APPROACH/ITEM TYPE:

The spiral testing booklet is designed to show one page to the student while the examiner reads instructions and cues from the other page.
Examiners record the errors on a summary sheet. This is used to pinpoint specific skills that need to be addressed during instruction.

The items include: reading sight words; making rhymes; reading words with an initial dipthong, with blends, with silent letters; reading a paragraph aloud to check on targeted words; reading, or listening to, short paragraphs aloud to

check comprehension via oral questions and responses. These paragraphs are based on adult experience stories.
Items become progressively more difficult.

NORMS: none, LVA suggests developing norms for each particular program.

COMMENTS:

This is a low level literacy level test which reflects current ideas of literacy test design, such as the inclusion of oral and written cues, and the option for listening or reading comprehension of the short paragraphs. The paragraph to be read aloud for targeted words, though, is written above a low literacy level.

This test, however, has no workplace orientation. The words and paragraphs are not job-related and no job-like documents (i.e., forms and schedules) are included.

TITLE: TABE (Tests of Adult Basic Education), Form E (easiest level)

PUBLISHER: CTB/McGraw-Hill

PURPOSE: to assess students' pre-instruction knowledge of reading, mathematics and language, to indicate areas of weakness in those skills, to assess achievement after instruction in those three skills

AUDIENCE: adults in ABE programs, also for adults with limited education and/or culturally disadvantaged backgrounds

TEST TYPE: WRITTEN

LEVELS: 4

FORMS: 6

TEST DESIGN: This test is based on the California Achievement Test (CAT) used in public schools.

paper and pencil test

all multiple choice responses

SKILLS ASSESSED: reading vocabulary, reading comprehension, mathematic computations, and mathematical concepts and applications (language mechanics, language expression, and spelling are found in the more difficult levels)

TESTING APPROACH/ITEM TYPE:

Vocabulary - utilizes a listening comprehension - sight word exercise and a synonym recognition exercise.

Comprehension - Has an alphabetizing exercise and comprehension questions for reading passages, tables of content and indexes.

Computations - checks basic arithmetic skills

Concepts and Applications - Concepts include telling time, reading thermometers and recognizing math symbols. Applications contain easy word problems to solve.

All items have multiple choice responses.

NORMS: No reliability information was provided in the manual. The manual claims the test has content validity because it is based on the CAT which has been shown to be valid.

COMMENTS:

The easiest level can be used with low level literacy students, but the content is not designed to test workplace literacy. None of the language nor tasks are job-related, except perhaps the reading comprehension exercises about tables of content and indexes.

Local Literacy Program Testing

To help determine the extent of the use of commercial tests, some local literacy programs in the Washington, D.C. area were contacted. These included literacy councils, local adult education programs and local vocational/job training programs. Most of the literacy councils use the Laubach system of instruction and extract exercises from the book lessons to create an assessment instrument for incoming students. The lessons are based on the phonetic approach to teaching reading and the content has little to do with an adult's daily life and/or work environment. These councils do not use commercial functional literacy tests.

Within the adult education programs, some use of commercial tests is made, but primarily any literacy assessment done is the result of a locally designed literacy test. The locally designed tests frequently check for left-right orientation, sight-word recognition, alphanumeric and sentence copying ability and short paragraph reading comprehension. Some teachers request brief student essays on an assigned topic. Some programs also perform oral interviews. The BVOPT, HELP and BEST are among the few commercial tests used. Since the majority of the adult students in the Washington metropolitan area is language minority, the use of these tests is understandable. HELP and BEST, as explained above, are not workplace oriented. BVOPT is; but it is only an oral test, not a literacy measure.

The vocational/job training programs also do not tend to utilize commercial tests. Some of the programs contacted perform no literacy testing. Others use their own instruments created from textbook exercises and teacher-made writing questions. BEST and Weschler I.Q. tests are used in a few of these programs, yet not on a regular basis with all students.

Although methods used by these contacts should not be generalized nationwide, they do seem to indicate that the commercial tests which assess basic skills are not commonly used in local literacy or adult education or vocational/job training programs. The fact that they are not particularly used in the job-related programs further supports the assertion that these tests are not appropriate for workplace literacy measurements.

Developing a Workplace Literacy Test

Envisioning the framework upon which to construct a workplace literacy test is relatively easy. It is the content and mode of delivery which is difficult. A common model used by most workplace literacy researchers for content incorporates information gathered through: 1) interviews with the employees and the employers, 2) analyses of job-related printed materials, 3) observations of the workplace in action, and 4) a review of the literature that already describes certain literacy requirements.

Several researchers have gone beyond these four procedures. Jacobs (1982) in her dairy plant project also photographed the written environment around the plant, i.e., the signs, warnings, gauges, etc., especially those with alphanumeric symbols, and she surveyed the employees through self-report questionnaires about their educational level and history. Latham & Parry (1979) explained that the content for the Sheffield Occupational Functional Reading Project pilot test used descriptions of training procedures from programs devoted to vocational training, as well as the four other procedures outlined above.

Determining the readability levels of job materials has also been used by several researchers. Moe et. al. (1980) evaluated the required reading materials (from 2000+ word samples) using the Dale-Chall Readability Formula and the Fry Readability Graph. They also recorded one hour oral samples at worksites to determine the speaking and listening requirements. They rated the speech for its formality and technical quality and also computer analyzed it for vocabulary and syntax structures. By reviewing the written and oral samples, they formed a key technical vocabulary list. They further organized the vocabulary to compose a highest frequency word list for each of the ten occupations they studied.

Sticht and his colleagues found readability indexes for Army manuals and analyzed the structural properties of the job reading materials before developing the JRRTs. Sticht and Caylor (1972) explained the importance of requiring the workers to show them the materials used on the job: "Job reading materials identified by interviewees may differ if the latter are permitted to simply state what they read, rather than being required to obtain and designate the reading materials they have used..." (p 49). Sticht and Mikulecky (1984) discussed the procedures for structured interviewing of the workers--ask them to describe the information they would look for to perform a task and then have them physically indicate the passages they would use in specific manuals.

Diehl (1978), in his model for assessing job literacy, included an analysis of attitudes, behaviors and strategies. For determining the literacy demands, Diehl suggested reviewing: the different kinds of materials for their reading and writing scope and depth; the amount of time actually spent reading on the job; and the availability of alternate information

sources. He also recognized that job training and retraining programs sometimes have greater literacy demands than the position necessitates and therefore the workers' performance in the program may not be totally indicative of their performance on the job. In order to reconcile this potential conflict, Diehl recommends an original model that first asks workers to relate:

1) their attitudes towards reading and towards their jobs;

2) their experience/background with reading; and

3) their experience on the job.

Then, through role plays and task completions, the workers' literacy strategies (i.e., how they cope with the reading required to complete certain tasks) are assessed.

Clearly, uncovering the appropriate content for a workplace test is a difficult and time-consuming process. As mentioned earlier, delivering that content through a suitable test design is also complicated. Many companies and businesses have chosen to use commercial basic skills and literacy tests because of their ready availability. However, the tests, such as those cited above, are inapplicable to the workplace for at least two reasons. First, pre-employment tests are difficult to read and often test at a reading level above a particular job's requirements. [B. Anderson (1981) cites research by Schultz (1975) and Hunt & Lindley (1977).] Second, as *The Bottom Line*, in explaining certain problems inherent to workplace literacy program evaluation, states, "There are no common criteria for evaluating the performance of adult education programs in general. Improvements in reading based on the administration of a standardized reading test seem to have very little impact on job performance..." (p 37).

Before designing a workplace literacy test, several questions need to be answered. For instance, can one test be developed for a variety of work sites? Research has shown that certain skills and areas of knowledge do transfer across work domains, but a large number of the required skills for one specific position are not necessarily required in others. Hence, a test that addresses the general, transferable skills may be little better than what is currently available. Likewise, a test that addresses the specific requirements of one job would be unsatisfactory for an unrelated position. The idea that one workplace literacy test can be created to assess all workers is perhaps unrealistic.

Since the tests need to be job specific, the possibility of clustering related jobs has merit. If several tests were designed around clusters (e.g., clerical, machinery-related, services-oriented, etc.), the tests would be both more face-valid and more valuable. A feasible design could include one part, which would measure those general, transferable, problem-solving skills, and remain the same for each test; and a second part, which would assess the more job specific skills, and change according to the occupation/cluster.

Another question that needs to be faced is whether a paper and pencil test is sufficient to measure non-paper and pencil workplace literacy tasks. Most current tests do not require simulated job performance using work-related materials. Simulation would allow the test to assess more closely a (potential) employees' ability to comprehend the material and perform the actual job tasks. Furthermore, having the opportunity to use manipulatives often helps decrease the test anxiety experienced by some individuals.

An additional question centers around the use of technology in the testing situation. Computer adaptive testing has vastly improved the ability to assess the depth of an individual's knowledge. For example, a correct response given by the examinee can induce the computer to bypass a certain number of similar questions once knowledge of that particular subject has been demonstrated. Or, an incorrect response can trigger the computer to present a question of similar content but with a different format. If several incorrect responses are given despite a variety of question phrases, then the test administrator can more readily assume a lack of knowledge on the examinee's part. The computer can hold all the information, record test scores, determine increments of change with later tests during a training program, for instance, and in general, keep track of an individual's progress. Should computer adaptive testing become the norm for workplace skills assessment rather than the exception?

It should be noted that interactive videodiscs afford even greater options than computers. They have the same capabilities as computers with the tracking and controlling of questions, but their advantage is the close simulation of real-life, problem solving tasks. Auditory and visual cues are presented to an examinee or student, as well as the traditional written ones. Interactive videodiscs also provide an excellent opportunity to assess more individuals, because they permit both visual and oral/aural learners to participate. Moreover, most individuals learn through more than one mode (although they usually have a preference) and these discs offer several modes of testing and learning. Manipulation can also be done in videodisc by moving or connecting graphic images, instead of handling

actual parts. An example of videodisc's effectiveness as a testing tool can be found in a widely-used disc which tests responses to medical emergencies by hospital emergency room personnel. Confronted with a bloody or comatose patient on the video screen, the doctors, nurses and attendants need to administer the right medicines and perform the right procedures, or the patient dies.

While videodisc can ameliorate some problems concerning learning styles, it cannot be the only answer. Many people do learn (and test) well through written information; but extensive research has shown that some people learn best through oral/aural communication, while others learn best through visual demonstrations and hands-on practice. The current focus on teams in the workplace and cooperative education in our classrooms should be explored for its application to testing. Using information sources, such as resource materials or partners, while taking a test, actually better correlates with the realities of the workplace. How can such strategies be incorporated into workplace tests?

This question as well as the others raised within this paper need to be discussed, and possible solutions need to be proposed, before a test or tests of workplace literacy should be developed. A new initiative could begin with a conference on workplace literacy that gathers testing experts, and management, union and training personnel as particpants. This would be a logical first step towards addressing the problems surrounding the assessment of workplace literacy. During that conference, feasible test designs, item content and potential uses would be explored. Since at present there is no consensus about the composition of a workplace literacy test, such a conference would provide a forum for debate and subsequently, proposed guidelines. These guidelines, having the

approval of the major players in the workplace literacy field, could then engender the development of integrative assessment measures, acceptable for field-testing and revisions.

Summary

This paper has been concerned with the assessment of workplace literacy. It has discussed various former and current definitions of literacy, the feasibility of using commercial tests to help assess the literacy skills of the workers, and the need to determine the job literacy requirements for employment positions. It has also evaluated the tests most commonly used in the workplace .

The paper has shown that definitions of literacy, and workplace literacy in particular, have changed over time and are continuing to change. At present, there is no single, widely-accepted definition of literacy. Moreover, reading grade levels which are frequently used to describe either demands or competence are not suitable for workplace literacy.

The paper has explained that most tests currently in use are inappropriate for the work environment, primarily because they do not address job-related vocabulary, nor tasks. Very few of the tests described are indicative of the actual demands placed upon a worker in the workplace. Most tests are of the paper and pencil variety, rather than performance-based. Very few contain an oral/aural component, although several studies have shown that managers consider "following oral instructions" and "expressing ideas and problems" serious deficiencies among their employees' abilities. Another gap not filled by many tests is the presentation of problems requiring the examinee to interpret and analyze graphs, tables, charts, etc.--skills frequently needed on the job.

The question of the possibility of developing an appropriate workplace literacy test has also been raised in this paper. It was strongly recommended that a needs analysis of the actual literacy requirements for each job be undertaken before a test of workplace literacy is designed. Furthermore, it was suggested that the advent of new technology in the teaching and testing fields should be incorporated into any workplace test. The capabilities of computers and interactive videodiscs permit more flexibilty, greater tracking of student progress, and better curriculum planning (i.e., for a training or retraining program) than any standardized test now affords. Interactive videodiscs in particular allow the inclusion of a wide variety of question types, ranging from the traditional format (with multiple-choice, true/false, matching, etc.) to a manipulative approach whereby the examinee can move or connect graphic images to simulate the activities of a workplace.

It has been concluded that one single workplace test would be unlikely to offer much face validity. A possible solution would be tests with two sections: one offered to all professions and addressing common concerns like reading indexes and tables of contents, filling out forms, writing business correspondence; the another varied according to similar occupations (i.e., clerical, service-oriented, machine-related, etc.) and focusing questions on the particular skills and literacy requirements needed for success in those job clusters.

The lack of consensus regarding the design, content and use of a workplace literacy test indicates the need for a national conference, bringing

together experts from the testing field and personnel from management, unions, and training programs. The conference could result in proposed guidelines for creating a test or tests which would have the approval of these domains.

REFERENCES

Anderson, B. (1981, June). *Guide to Adult Functional Literacy Assessment Using Existing Tests.* Portland, OR: Northwest Regional Educational Laboratory.

Anderson, N. (1987). Henderson-Moriarity ESL Placement Test (HELP). In J. C. Alderson, K. J. Krahnke, & C. W. Stansfield (Eds.), *Reviews of English Language Proficiency Tests* (pp. 35-37). Washington, DC: TESOL.

Blocklyn, P. L. (1988, February). Preemployment Testing. *Personnel 65(2),* 66-68.

Broussard, R. L. (1985, November). *An Evaluation of the Cultural Bias of the Adult Performance Level Assessment.* Paper presented at the annual meeting of the Mid-South Educational Research Association. Biloxi, MS.

The Bottom Line: Basic Skills in the Workplace. (1988). Washington, DC: US Department of Education and US Department of Labor.

Cervaro, R. M. (1981). Assessment in Adult Basic Education: A Comparison of APL and GED Tests. *Adult Education 31(2),* 67-84.

Cronbach, L. J. (1978). USES Basic Occupational Literacy Test. In O. K. Buros (Ed.), *The Eighth Mental Measurements Yearbook,1* (pp. 673-674). Highland Park, NJ: The Gryphon Press.

Delker, P. (1983, May). Defining Adult Functional Literacy. *Functional Literacy and the Workplace. Proceedings of a National Invitational Conference* (pp. 31-33). Washington, DC: American Council of Life Insurance.

Diehl, W. (1978, April). *The Variable and Symbolic Natures of Functional Literacy.* MA Thesis. Indiana University, Bloomington, IN.

Harman, D. (1985). *Turning Illiteracy Around: An Agenda for National Action, Working Paper 2.* New York: Business Council for Effective Learning.

Harris, L. and Associates. (1971, August). *The National Reading Difficulty Index: A Study of Functional Reading Ability in the US.* Washington, DC:

National Reading Center.

Heinemann, S. T. (1979, December). Can Job-Related Tasks Be Used to Diagnose Secretaries' Reading and Writing Skills? In J. R. Binkley (Ed.), *Journal of Reading* (pp. 239-243). Newark, DE: International Reading Association.

Henry, J. F. (1983, May). Expectations of the Workplace: Basic Skills in the US WorkForce. *Functional Literacy and the Workplace. Proceedings of a National Invitational Conference* (pp. 21-24). Washington, DC: American Council of Life Insurance.

Jacob, E. (1982, June). *Literacy on the Job.* Final Report of the Ethnographic Component of the Industrial Literacy Project. Washington, DC: Center for Applied Linguistics.

Johnston, W. B. & Packer, A. H. (1987). *Workforce 2000: Work and Workers for the Twenty-First Century.* Indianapolis, IN: Hudson Institute.

Kirsch, I. & Jungeblatt, A. (1986). *Literacy: Profiles of America's Young Adults.* Final Report (Report No. 16-PL-01). Princeton, NJ: National Assessment of Educational Progress.

Latham, W. & Parry, O. (1979, September). *A Criterion-Referenced Test of Occupational Functional Reading Ability.* Paper presented at the anual meeting of the Reading Association of Ireland. Dublin, Ireland.

Leibert, R. E. (1973, January) *Establishing Criteria for Instructional Placement on the Basis of Performance on the Adult Informal Reading test and a Standardized Test.* University of Missouri-Kansas City [OEG-0-71-4399 (324)].

Messerschmidt, D. (1987). Bilingual Vocational Oral Placement Test (BVOPT). In J. C. Alderson, K. J. Krahnke, & C. W. Stansfield (Eds.), *Reviews of English Language Proficiency Tests* (pp. 17-18). Washington, DC: TESOL.

Mikulecky, L., Ehlinger, J., & Meenan, A. (1987). *Training for Job Literacy Demands: What Research Applies to Practice.* University Park, PA: Penn State University, College of Education.

Mikulecky, L. & Diehl, W. (1979, July). *Literacy Requirements in Business and Industry.* Paper for the US Department of Housing, Education and

Welfare. Bloomington, IN: Reading Research Center at the Indiana University.

Moe, A. J., Rush, R. T., & Storlie, R. L. (1980, January). *The Literacy Requirements of an Industrial Mechanic on the Job and in a Vocational Training Program.* Lafayette, IN: Department of Education at Purdue University.

Nafziger, D. H., Thompson, R. B., Hiscox, M. D., & Owen, T. R. (June 1975). *Tests of Functional Adult Literacy: An Evaluation of Currently Available Instruments.* Report for the US Office of Education (Contract No. 300-75-0098). Portland, OR: Northwest Regional Educational Laboratory.

Niebuhr, H. L., Jr. (1983, May). Perspectives on Literacy: Society's Stake in a Literate Population. *Functional Literacy and the Workplace. Proceedings of a National Invitational Conference* (pp. 17-20). Washington, DC: American Council of Life Insurance.

Northcutt, N. & Hickok, D. (1979). APL and ESL: Some Ideas for Implementation. In D. E. Bartley (Ed.), *The Adult Basic Education TESOL Handbook* (pp. 186-203). New York: Collier MacMillan International, Inc.

Rush, R. T., Moe, A. J., & Storlie, R. L. (1986). *Occupational Literacy Education.* Newark, DE: International Reading Association.

Sticht, T. G. & Mickulecky, L. (1984). *Job-Related Basic Skills: Cases and Conclusions.* Washington, DC: National Institute of Education.

Sticht, T. G. & Beck, L. J. (1976, June). *Development of an Experimental Literacy Assessment Battery.* Final Report. Alexandria, VA: Human Resources Research Organization.

Sticht, T. G. & Caylor, J. S. (1972). Development and Evaluation of Job Reading Task Tests. In L. E. Hefner (Ed.), *The Journal of Reading Behavior, 4(4),* 29-50.

Valentine, T. (1986, July). *Issues Central to the Definition of Adult Functional Literacy.* Paper presented for the Office of Higher Education and Adult Learning of the OERI of the US Department of Education (No. OERI-P-86-3014).

The Written Word. (1988, February-March). Lincoln, NE: Contact Literacy Center.